



## UWS Academic Portal

### **Efficient QoE-aware scheme for video quality switching operations in dynamic adaptive streaming**

Irondi, Iheanyi; Wang, Qi; Grecos, Christos; Alcaraz Calero, Jose M.; Casaseca, Pablo

*Published in:*

ACM Transactions on Multimedia Computing, Communications, and Applications

*DOI:*

[10.1145/3269494](https://doi.org/10.1145/3269494)

E-pub ahead of print: 01/02/2019

*Document Version*

Peer reviewed version

[Link to publication on the UWS Academic Portal](#)

*Citation for published version (APA):*

Irondi, I., Wang, Q., Grecos, C., Alcaraz Calero, J. M., & Casaseca, P. (2019). Efficient QoE-aware scheme for video quality switching operations in dynamic adaptive streaming. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 15(1), [17]. <https://doi.org/10.1145/3269494>

#### **General rights**

Copyright and moral rights for the publications made accessible in the UWS Academic Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

#### **Take down policy**

If you believe that this document breaches copyright please contact [pure@uws.ac.uk](mailto:pure@uws.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

© ACM, 2018. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in ACM Transactions on Multimedia Computing, Communications, and Applications, {VOL#, ISS#, (DATE)}  
<http://doi.acm.org/10.1145/3269494>

[Details to be complete at time of full publication]

# Efficient QoE-Aware Scheme for Video Quality Switching Operations in Dynamic Adaptive Streaming

IHEANYI IRONDI, and QI WANG, University of the West of Scotland

CHRISTOS GRECOS, Central Washington University

JOSE M. ALCARAZ CALERO, University of the West of Scotland

PABLO CASASECA-DE-LA-HIGUERA, University of the West of Scotland, and Universidad de Valladolid

Dynamic<sup>1</sup> Adaptive Streaming over HTTP (DASH) is a popular over-the-top video content distribution technique that adapts the streaming session according to the user's network condition typically in terms of downlink bandwidth. This video quality adaptation can be achieved by scaling the frame quality, spatial resolution or frame rate. Despite the flexibility on the video quality scaling methods, each of these quality scaling dimensions has varying effects on the Quality of Experience (QoE) for end users. Furthermore, in video streaming, the changes in motion over time along with the scaling method employed have an influence on QoE, hence the need to carefully tailor scaling methods to suit streaming applications and content type. In this work, we investigate an intelligent DASH approach for the latest video coding standard H.265 and propose a heuristic QoE-aware cost-efficient adaptation scheme that does not switch unnecessarily to the highest quality level but rather stays temporarily at an intermediate quality level in certain streaming scenarios. Such an approach achieves a comparable and consistent level of quality under impaired network conditions as commonly found in Internet and mobile networks whilst reducing bandwidth requirements and quality switching overhead. The rationale is based on our empirical experiments, which show that an increase in bitrate does not necessarily mean noticeable improvement in QoE. Furthermore, our work demonstrates that the Signal-to-Noise Ratio (SNR) and the spatial resolution scalability types are the best fit for our proposed algorithm. Finally, we demonstrate an innovative interaction between quality scaling methods and the polarity of switching operations. The proposed QoE-aware scheme is implemented and empirical results show that it is able to reduce bandwidth requirements by up to 41% whilst achieving equivalent QoE compared with a representative DASH reference implementation.

CCS Concepts: •**Computer systems organization** → Client-server architecture; •**Computing methodology** → Visual inspection

## KEYWORDS

QoE-aware, DASH, quality scaling dimension, adaptation algorithm

## ACM Reference format:

---

This work is in part supported by the European Commission Horizon 2020 5G-PPP Programs under Grant Agreement number H2020-ICT-2014-2/671672 (SELFNET: Framework for Self-Organized Network Management in Virtualized and Software Defined Networks), and H2020-ICT-2016-2/761913 (SliceNet: End-to-End Cognitive Network Slicing and Slice Management Framework in Virtualised Multi-Domain, Multi-Tenant 5G Networks). Thanks are also given to the to the UK Royal Society of Edinburgh and National Science Foundation of China for the funding associated to the project "Flood Detection and Monitoring using Hyperspectral Remote Sensing from Unmanned Aerial Vehicles".

Authors' addresses: I. Irondi, Q. Wang (Corresponding Author), J. Alcaraz-Calero, and P. Casaseca-de-la-Higuera, School of Engineering and Computing, University of the West of Scotland, Paisley PA1 2BE, United Kingdom; C. Grecos, Computer Science Department, Central Washington University, Ellensburg, WA 98926-7520, USA. P. Casaseca-de-la-Higuera is also with Laboratorio de Procesado de Imagen. ETSI Telecomunicación, Universidad de Valladolid. Valladolid, 47011, Spain.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

Copyright © ACM 2018 1551-6857/2018/MonthOfPublication - ArticleNumber \$15.00

<https://doi.org/10.1145/3269494>

Iheanyi Irondi et al. 2018. Efficient QoE-Aware Algorithm for switching Operations in Dynamic Adaptive Streaming. *ACM Trans. Multimedia Comput. Commun. Appl.* XX, X, Article XX (XXX 2018), 22 pages. DOI: 10.1145/3269494

## 1 INTRODUCTION

There has been a surge in the use of streaming technologies, multimedia services and applications in recent years. A recent forecast [1] has predicted that by 2019 the consumer Internet video will account for 80% of the total global traffic and long-form (more than 10 minutes of duration) video will be 72.9% of the Internet video traffic. A similar trend has been predicted for mobile networks such as the emerging Fifth-Generation mobile networks (5G). Considering the limitations in bandwidth and the error-prone nature of the Internet and mobile networks, the Dynamic Adaptive Streaming over HTTP (DASH) technique allows video streaming sessions to be adapted to the user's network condition typically in terms of downlink bandwidth through switching among predefined video segments of different bitrate levels. Video can be adapted by scaling different dimensions such as the frame quality, frame rate, frame resolution etc. to achieve different video bitrate levels in DASH. Moreover, different segment sizes have varying effects on users' Quality of Experience (QoE). Furthermore, the different quality scaling methods used in DASH introduce their own individual challenges such as visual artifacts that affect QoE.

In this paper, we investigate optimization schemes for DASH-based adaptive video streaming, focusing on improving transmission efficiency (bandwidth saving) without compromising users' QoE as the design goal. We follow an empirical approach and conduct three different subjective evaluations including initial QoE measurements to motivate and inform the subsequent design, hypothesis and evaluation for a novel optimization algorithm, and further QoE validation of the algorithm. We also perform a fourth experiment to evaluate bandwidth savings using the proposed algorithm. A summary of these four experiments is shown in Table 1.

Based on these experiments, we evaluate how users perceive the interactions between different quality scaling techniques and various switching patterns/scenarios as the motion complexity within the video sequence changes over time. Our experimental findings show that an increase in bitrate level does not necessarily improve the QoE and we also demonstrate how the amplitude of switching influences the QoE. We then propose a heuristic QoE-aware cost-efficient framework that does not switch quickly to the highest bitrate under certain circumstances, but rather stays temporarily at an intermediate bitrate level to reduce bandwidth requirements whilst achieving a comparable QoE level with the highest bitrate level. It is worth highlighting that we employ the latest video codec standard H.265 (also known as High Efficiency Video Coding or HEVC) [2] and thus the obtained results are more relevant to the new-generation video streaming applications such as those envisioned for 5G mobile networks in line with the industry's vision [3].

The contributions of our proposed system are fivefold. First, the increased period of switching minimizes unnecessary switching, which causes flicker in very unstable network environments, thereby improving the QoE. Our proposed system isolates short-term throughput variations (as a result of TCP congestion control) from throughput congestions (as a result of persistent throughput variations). Secondly, the proposed scheme saves bandwidth while achieving a comparable level of quality with an intermediate bitrate, chosen temporarily as the optimal bitrate in fluctuating network conditions. Thirdly, our proposed scheme minimizes the amplitude of switching operations in fluctuating network conditions, which also improves the QoE. The fourth contribution is the heuristic reduction of inaccurate bandwidth estimations caused by short segment sizes, where short-term network capacity measurements lead to non-optimum quality adaptations. Finally, the achieved bandwidth savings also reduce power consumption with the consequent benefits for mobile devices, where battery resource is an ongoing issue coupled with processing capacity.

The remainder of this paper is structured as follows. Section 2 provides an overview of adaptive video streaming and reviews the current state of the art. The initial empirical QoE experiments that have inspired our subsequent proposal are discussed in the section 3. Section 4 presents our novel adaptation system to achieve QoE-aware, cost-efficient DASH streaming, and analyzes the experimental results. Finally, the paper is concluded in section 5.

**Table 1. Summary of Performed Experiments**

Experiment	Type	Description	Stimuli	Sessions
1	Subjective Evaluation	Initial QoE experiments analyzing different quality switching patterns and various switching patterns to inform the design.	4	20
2	Subjective Evaluation	Further evaluation/hypothesis for potential bandwidth savings at comparable QoE with an intermediate quality level.	4	9
3	Bandwidth Savings	Demonstration of achieved bandwidth savings with the proposed algorithm.	1	10
4	Subjective Evaluation	Demonstration of achieved comparable level of QoE with proposed algorithm.	4	6

## 2 BACKGROUND AND RELATED WORK

### 2.1 Overview of Adaptive Video Streaming over HTTP

MPEG-DASH [4] is a standard ISO/IEC MPEG adaptive streaming technology for multimedia content delivery using existing web servers. This technique is designed to address the shortcomings of the progressive download and the conventional Real-time Transport Protocol (RTP) over User Datagram Protocol (UDP) based streaming, and it is suited for cases where there is constant variation in bandwidth or other changing network conditions. DASH runs on the pervasive Hypertext Transfer Protocol (HTTP) over the Transmission Control Protocol (TCP), and thus it does not need deployment of specialized web servers and leverages the same existing Internet infrastructures including caches, firewalls and Content Delivery Networks (CDNs). Using the HTTP stateless protocol, DASH can easily traverse firewalls, thus making it more attractive when compared with traditional streaming protocols such as RTP with the Real-time Streaming Protocol (RTSP). Meanwhile, it is noted that a new reliable and secure experimental protocol called Quick UDP Internet Connection (QUIC) has recently been proposed by Google to reduce transport and connection latency [5]. Nevertheless, DASH has gained global momentum in light of the advantageous benefits as aforementioned.

In DASH, a multimedia file is encoded into different bitrate representations, segmented into chunks, and then delivered to the client using HTTP from a standard web server. The segmented multimedia files are stored together with the Media Presentation Description (MPD), an eXtensible Markup Language (XML) file that describes the various media segments stored on the server. It is noted that there is no standard segment size in DASH, although Microsoft and Apple recommend 2s [6] and 10s [7] durations respectively in their own proprietary solutions, which motivated the DASH standard.

### 2.2 Metadata-Based Adaptive Video Streaming Techniques

In DASH, the streaming session is controlled by the client and the mechanism for client adaptation to available bandwidth is beyond the scope of the DASH standard where different designers employ different clients, resulting in varying streaming experience and hence varying effects on the QoE. As the streaming session is controlled by the client, in existing work a lot of emphasis is on improving the client side in order to improve the QoE. For instance, Akhshabi et al. [8] evaluated the rate adaptation algorithms of Adobe OSMF, Microsoft smooth streaming and Netflix. Their findings reveal that the technology is still new and that there is a need for further improvement in the existing adaptation algorithms. Li et al. [9] studied the effect of multiple clients competing for the same available bandwidth and demonstrated the difficulty that clients face in correctly predicting their own fair share of bandwidth. The authors then proposed an adaptive video streaming rate adaptation algorithm that minimizes the instability of bitrate selection without increasing the chances of buffer underflow. Through experimental studies, Huang et al. [10] realize that capacity estimation in environments with highly variable throughput is challenging and inaccurate. They proposed an adaptation algorithm that

starts by selecting the video representation based on the current buffer occupancy and then employs capacity estimation when needed.

Moreover, performance can also be improved at the server side. For example, Gürses et al. [11] proposed a system comprising an input buffer at the server side, with TCP congestion control mechanism for efficient streaming. Their proposed scheme selectively discards frames with low priority that would have compromised the successful playout of frames with high priority. Van der Hooft et al. [12] proposed an HTTP/2 push-based approach for HTTP adaptive video streaming, which reduces the Round-Trip Time (RTT) of shorter segment sizes by pushing on segments of video content to the client once the MPD request is received. However, the use of smaller segment size comes at the cost of increased overhead.

### 2.3 QoE-aware and Video Quality Adaptation Techniques

Most of the existing adaptation algorithms in DASH are based on throughput measurement and buffer measurement rather than QoE driven. Furthermore, the existing quality assessment metrics are designed for traditional streaming and cannot capture the issues created by adaptive video streaming. To design efficient adaptation algorithms and quality assessment metrics, there is a need to understand the key factors affecting the QoE in adaptive video streaming. In DASH, different bitrate representations of the same video content are segmented into chunks at Group Of Pictures (GOP) boundaries to allow seamless adaptation of the streaming session according to the user's bandwidth condition. There is no defined duration for the segments in the DASH standard and different implementations have their own segment durations with varying effects on QoE. Bitrate switching in DASH involves adapting the streaming session according to the user's bandwidth condition. Most adaptive streaming players by default start streaming from the lowest bitrate and switch to higher bitrates if the network conditions permit. The players can decide to switch to lower bitrate representations once the network conditions deteriorate, in order to minimize stalling during playback. The frequency of switching events between different video quality levels and the rate of bitrate levels (amplitude of switching) while switching at any point, both have effect on the QoE in a streaming session. Depending on the number of bitrate representations to choose from, switching can be abrupt or step-wise.

Frequent switching of quality should be avoided to allow users to get accustomed to a given video quality [13], and if switching cannot be avoided, the amplitude of switching should be minimized [14], [15]. Switching down to a lower quality is more noticeable compared with switching up to a higher quality [16]. Rodríguez et al. [17] attempted to capture the relationship between type, frequency and temporal location of switching events and the perceived quality, and showed that as the frequency of switching increased, the Mean Opinion Score (MOS) values significantly dropped and frequent bitrate level switching had a high impact on QoE. Spotting oscillations in video quality is challenging with persistent scene changes [18] and change in quality is perceptible if users can predict the next scene. Persistent quality or nearly persistent quality is often preferred to frequently changing quality, even in cases where quality is higher than the mean quality [18]. In addition, a negative impact on QoE occurs by an increase in the initial delay, or in the frequency and duration of stalling [19-21].

Adaptive streaming involves adapting the streaming session according to the users' bandwidth condition. This adaptation necessarily involves choosing video segments from different bitrate levels of the same video. The different bitrate levels can be achieved by varying any of the scaling dimensions such as frame quality, spatial resolution or frame rate during encoding. Layer switching with any of the methods introduces visual artefacts referred to as noise flicker, blur flicker and motion flicker respectively [22]. For frame rate adaptation, there is always lack of fluency of video motion at lower frame rates [22]. Different types of video quality levels have different effects on QoE and therefore the adaptation techniques should be considered to achieve optimal QoE [17]. Tavakoli et al. [23] studied the effect of various adaptation strategies on QoE. Their results showed the huge impact of content type and its spatio-temporal information had on QoE. Huynh-Thu and Ghanbari [21, 24, 25] showed that at a given frame rate, the QoE of low motion contents was not always better than that high motion content and also that head and shoulder sequences should be considered as a particular class of videos, since the level of motion is not the only factor influencing perceived quality. They also showed that jitter was worse than jerkiness and that the perceived impact of jitter strongly depended on the content type.

**Table 2. Absolute Category Rating (ACR)**

Value	Scale
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Despite the above advances in DASH, there are still a number of challenges to be addressed and gaps to be filled. In particular, studies on practical and efficient QoE-aware DASH techniques utilizing the latest video coding standard H.265/HEVC have still not been sufficiently investigated. Our proposed algorithm in this paper minimizes the amplitude and frequency of switching in fluctuating networking conditions, which improves the QoE. The paper also highlights the impact of various quality adaptation techniques on QoE, especially the negative impact of explicitly scaling down the quality with the frame rate scaling technique because of the observed jerkiness.

### 3 SUBJECTIVE EVALUATION OF QUALITY SCALING DIMENSIONS

#### 3.1 Experiment design and Evaluation Methodology

This section presents preliminary experiments to investigate the perceived impact of different switching techniques on the QoE. The review of the state of the art in section 2 allowed us to identify the important factors enabling the novel contributions of this paper. The following subsections summarize this study.

##### 3.1.1 Test Method.

The rating method used for this subjective evaluation was the Absolute Category Rating (ACR) method as described in ITU [26]. In this method, the video clips were randomly selected and presented to the test subjects one at a time while different network impairments were added by the NETEM box placed in between the Client computer and the webserver as the video is presented. After each video presentation, the subjects then rated the videos independently using the ACR scale as shown in Table 2. Training on the correct use of the scale was provided before the experiments. The ACR scale in Table 2 ranges from 1 to 5, with the value of 5 representing the highest possible rating or the highest quality. The streaming sessions for the experiment involved the presentation of individual 60s videos, subjected to different streaming scenarios dependent on the study factors. The reference video was not disclosed to the test subjects. The 60s duration is considered sufficient to demonstrate the factors under investigation and is also longer than the duration often used in the literature. The subjects were also given frequent breaks within and among experiments. The test scores were collected and collated after the tests, checked for outliers and then a MOS was calculated. Finally, the MOS values were plotted graphically for final analysis.

##### 3.1.2 Test Conditions and Materials.

The experiment involves investigating the performance of adaptive video streaming based on the scaling dimensions of HEVC, where different perceived quality levels can be achieved by varying the frame quality through the Quantization Parameter (QP) parameter, the frame rate, and the spatial resolution.

The selection of video sequences for this experiment was based on the Spatial Index/Temporal Index (SI/TI) as per [27]. Each video was encoded with the HM reference software to obtain three bitrate representations or quality levels for each of the frame quality, frame rate, and spatial resolution scaling dimensions respectively. We then chose and adapted the smooth streaming pattern as motivated by Tavakoli et al. [28]. Four video sequences (“Spread”, “Bigbuck”, “Car” and “Tractor”) were selected for this experiment to illustrate the different complexities within different video sequences. Scene traits that can uniquely interact with the user’s perception or with the codec were also one of the main criteria in selecting these sequences.

Table 3. Video Sequences QoE Evaluation

Video Clip	Frame Rate	QP	Spatial Resolution	SI	TI	Description
Bigbuck Bunny	24, 12, 8	25, 30, 35	1980x1080 1280x720 832x480	51	6	Animated science fiction, medium motion with two main characters exploring the twisted and dark complex wires, cogs and gears of a giant machine [29].
Spread	24, 12, 8	25, 30, 35	1980x1080 1280x720 832x480	38	10	Short movie, high motion, rapid scene changes, lots of detail, multiple talking heads, group of people in a corporate environment having meetings, lunch and getting on an elevator.
Car	24, 12, 8	25, 30, 35	1980x1080 1280x720 832x480	27	8	Surveillance, high motion, fixed camera, blurred background, with in-focus foreground, cars travelling on a motorway.
Tractor	24, 12, 8	25, 30, 35	1980x1080 1280x720 832x480	70	16	Amateur filming, relatively still motion, no scene changes, moving camera, filming a tractor on an open field.

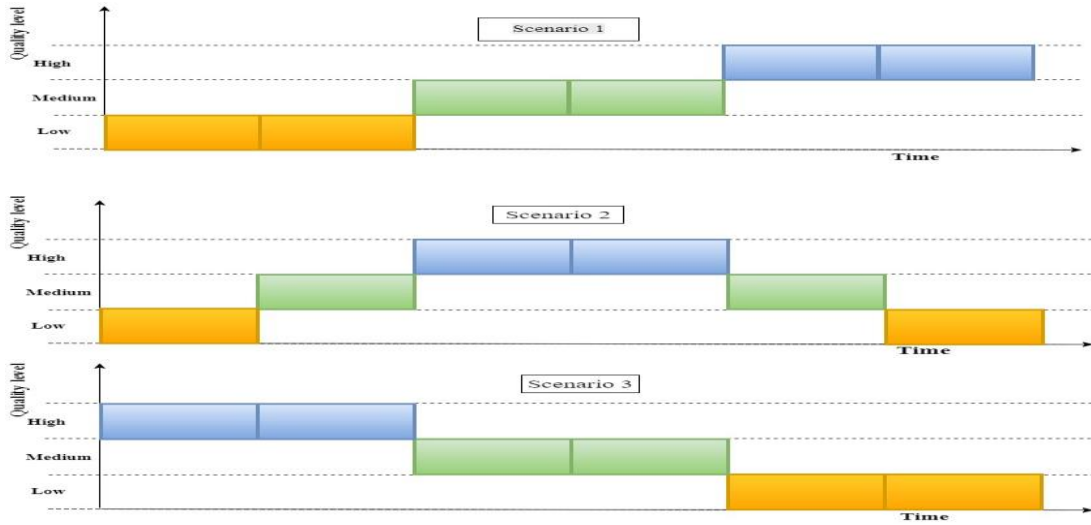


Fig. 1. Illustration of basic switching operations.



**Table 4. Quality Level Parameters for HEVC Scaling Dimensions**

	Low	Medium	High
Frame Quality (QP value)	35	30	25
Spatial Resolution (SR, width x height)	1920x1080 (1080p)	1280x720 (720p)	832x480(480p)
Frame Rate (FR, fps)	8	12	24

**Table 5. Streaming Scenarios**

Code	Resolution	Quality switching	Changing Dimensions	Scenario (Figure 1)
Z_480	480p	No switching, only streaming at high QP and high FR	N/A	
Z_720	720p	No switching, only streaming at high QP and high FR	N/A	
Z_1080	1080p	No switching, only streaming at high QP and high FR	N/A	
LMH_480	480p	Low-Medium-High	QP or FR	1
LMH_720	720p	Low-Medium-High	QP or FR	1
LMH_1080	1080p	Low-Medium-High	QP or FR	1
LHL_480	480p	Low-High-Low	QP or FR	2
LHL_720	720p	Low-High-Low	QP or FR	2
LHL_1080	1080p	Low-High-Low	QP or FR	2
HML_480	480p	High-Medium-Low	FR	3
HML_720	720p	High-Medium-Low	FR	3
HML_1080	1080p	High-Medium-Low	FR	3
LMH_SP	Varying	Low-Medium-High	SR	1
LHL_SP	Varying	Low-High-Low	SR	2

These videos were then encoded into three quality levels each (by varying the respective parameters) using the configurations shown in Table 3 for frame quality (based on the QP parameter), spatial resolution and frame-rate adaptation. Table 4 also shows the selected values for “high”, “medium”, and “low” quality for each of the dimensions. In order to realize 3 different quality levels, it should be noted during encoding for each individual resolution (480p, 720p, and 1080p); the QP is varied while the frame rate and resolution are kept constant for frame quality adaptation, the frame-rate is varied while the QP and resolution are kept constant for frame-rate adaptation. Then the resolution is varied while frame-rate and QP are kept constant for spatial resolution adaptation.

We selected a number of streaming scenarios where quality was changed using different patterns for each dimension. These include from no switching at all to a variety of transitions (e.g. Low-High-Low, High-Medium-Low, etc.), where each quality level had a total time of 60s (2s x 30 segments after encoding and segmentation). The streaming scenarios are described in Table 5 and further illustrated in Figure 1 (note that for scenario 2, the transition Low-High-Low involves intermediate medium levels). They have been carefully chosen to capture the primary switching styles that may be encountered in the course of real time video streaming and thus provide insights to other switching patterns later adopted. Table 5 also shows which scenarios have been used for the

assessment of scaling in each dimension. High-Medium-Low transitions have only been included in frame rate scaling to specifically assess the impact of jerkiness associated to downscaling frame rate.

According to Table 5, Z\_480, Z\_720, Z\_1080 denote streaming at highest quality level without switching for the 480p, 720p and 1080p resolutions respectively. LMH\_X denotes (Where X can be 480, 720, 1080 resolutions) switching from Low to Medium to High quality for different adaptation strategies. LHL\_X denotes (Where X can be 480, 720, 1080 resolutions) switching from Low to High and back to Low quality for different adaptation strategies. HML\_X (Where X can be 480, 720, 1080 resolutions) switching from High to Medium and back to Low quality for the frame-rate adaptation strategy. Finally, LMH\_SP denotes switching from 480p to 720p to 1080p and LHL\_SP denotes switching from 480p to 720p to 1080p and back to 480p.

### 3.1.3 Participants and test environments.

A total of 15 test subjects comprising 10 males and 5 females were involved in the experiment. Their ages ranged from 25 to 50 years old and they had no knowledge of video quality evaluation. All the volunteers were screened for normal visual acuity and normal colour vision as recommended by ITU [26] and ITU [27]. Furthermore, no health issues were reported by the subjects. The subjective evaluation was carried out in a controlled environment as per [26] guidelines. A 22-inch monitor was used for presenting the video clips to the subjects, who were asked to rate the videos immediately after viewing.

## 3.2 Results and Discussion

It should be noted that ideal streaming in our analysis refers to streaming at the highest quality level without switching (zero switching). The error bar used in our results is 95% confidence interval. MOS in the results and analysis refers to Mean Opinion Score (MOS) and refers to average score for each investigation. In addition, the Mann-Whitney U test was used as in [30, 31] for statistical significance pairwise in this experiment where the MOS is based on an ordinal scale.

### 3.2.1 Adaptation through frame quality scaling.

Figure 2 shows MOS values and error bars for the adaptation based on QP parameter frame quality scaling. Different streaming resolutions of 480p, 720p and 1080p were examined. The following main conclusions can be inferred from the figures: The “spread” video sequence has a higher perceived quality when there is no switching compared to other videos. Switching using the Low-Medium-High sequence involves an improvement in QoE for most of the video sequences except the “spread” video where users had better perceived picture quality with zero switching. In addition, the level of improvement differs for different resolutions. This suggests that the content type and the particular video quality level under investigation affect QoE when the quality is being switched.

MOS values decrease with the transition Low-High-Low, meaning that the users did not perceive any improvement in this transition, despite the interim switching to the highest quality and back to the lowest quality. This clearly shows that the QoE is affected with reduction in the quality of the video streaming session, even though a higher quality was briefly streamed. This is in line with the known recency effect where users tend to recall the perceived quality based on the end of the streaming session.

Using the Mann-Whitney test, we found that there are significant differences between L-M-H and the Low-High-Low transitions (Mann-Whitney  $p = 0.00001$ , level of significance = 5% for the 480p “spread” video). There are no significant improvements in QoE when switching to higher quality with 1080p resolution video for the “bigbuck” and “Car” clip (Mann-Whitney  $p = 0.77182$ , level of significance = 5% and  $p = 0.98404$ , level of significance = 5% respectively) compared to lower resolutions, while the “spread” and “tractor” clip did not show any improvement in QoE. This observation means that increasing the bitrate at that higher resolution might not necessarily mean an improvement in QoE. This in turn suggests that unnecessarily switching to a higher bitrate can thus be avoided, resulting in reduction of the thus encoding overhead. This evidence will help in formulating a hypothesis for proposing a new adaptation algorithm in section 4.

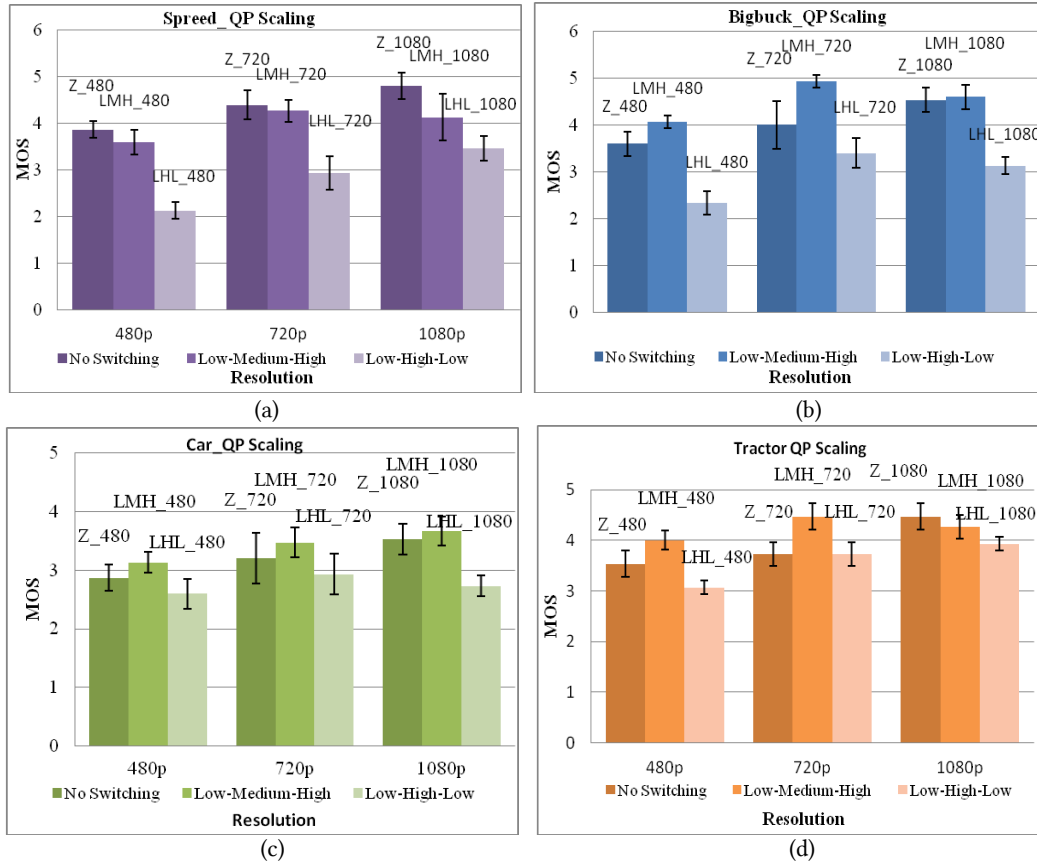


Fig. 2. MOS for QP quality scaling: (a) Spread clip, (b) Bigbuck clip, (c) Car clip, (d) Tractor clip.

### 3.2.2 Adaptation through Frame rate quality scaling.

Figure 3 shows the equivalent to Figure 2 for the adaption through frame rate scaling in various streaming scenarios. The following main conclusions can be extracted. There is a significant decrease in QoE at all resolutions when switching in Low-Medium-High (LMH) scenarios is performed compared with the highest quality (only streaming) except for the “tractor” sequence, which slightly increased for 480p and decreased for 720p and 1080p. This may originate from the fact that the tractor sequence has a much higher SI/TI values than other sequences; other possible reasons for this different behaviour might be that the tractor sequence took a longer time to encode than other sequences, and it also had a camera motion.

The results in Figure 3 also show that the video quality as perceived by human users is significantly impaired by frame rate quality switching, even at lower QP values. This is likely to be the result of the sustained stalling observed when the video is switched to a higher bitrate. A similar trend was observed when the streaming session was switched from the lowest quality to the highest quality and then back to the lowest quality. We additionally assessed High-Medium-Low (HML) scenarios to further investigate how QoE perception changes in scenarios where quality is reduced, compared to those in which, quality is improved. To do so, we had to explicitly start streaming from the highest quality and then smoothly switch to the lowest quality as compared against the conventional way of the streaming, which starts from the lowest quality and then progressively moves to higher qualities through switching. From the figures, we can see that the MOS values are lower for HML scenarios compared to LMH or LHL. This highlights some significant observations, namely that even though the test subjects were not impressed with switching up the quality level using the frame rate scaling

technique (perceived stalling) and that they were more dissatisfied with explicitly scaling down the quality with the frame rate scaling technique because of the observed jerkiness. These observations indicate that jerkiness and stalling have different effects on the perceived picture and need to be investigated further.

This shows that the video quality as perceived by human users is significantly impaired by frame rate quality switching, even at lower QP values. This is likely to be the result of the sustained stalling observed when the video is switched to a higher bitrate. A similar trend was observed when the streaming session was switched from the lowest quality to the highest quality and then back to the lowest quality. We additionally assessed High-Medium-Low (HML) scenarios to further investigate how QoE perception changes in scenarios where quality is reduced, compared to those in which, quality is improved. To do so, we had to explicitly start streaming from the highest quality and then smoothly switch to the lowest quality as compared against the conventional way of the streaming, which starts from the lowest quality and then progressively moves to higher qualities through switching. From the figures, we can see that the MOS values are lower for HML scenarios compared to LMH or LHL. This highlights some significant observations, namely that even though the test subjects were not impressed with switching up the quality level using the frame rate scaling technique (perceived stalling) and that they were more dissatisfied with explicitly scaling down the quality with the frame rate scaling technique because of the observed jerkiness. These observations indicate that jerkiness and stalling have different effects on the perceived picture and need to be investigated further.

### 3.2.3 Adaptation through spatial resolution quality scaling.

Figure 4 shows the MOS values for rating quality in spatial resolution of maximum 1080p with different switching patterns. The MOS values in the figures show that there is a decrease in quality perception when there is quality switching from the ideal streaming conditions at 1080p. It is worth noting that the perceived quality decrease is more relevant for the LHL pattern, following the same trend as in the QP and frame rates dimensions. The error bars generally overlapped for the streaming at maximum quality and switching from lowest to highest bitrate level (LMH\_SP), which suggests that the difference may or may not be statistically significant except for the LHL\_SP (“bigbuck”, “car” and “tractor” sequences), whose error bars do not overlap, suggesting that the difference is significant. Despite the difficulty in showing statistical significance, we can still make the following observations: there is a decrease in perceived quality as the video quality is being switched, the reduction in QoE is not as severe in the spatial scaling dimension as compared with the frame rate scaling dimensions and users had better QoE with the “spread” sequence in the spatial scaling dimension as compared with other sequences and other scaling dimensions.

Overall, there might not be any significant difference between the LMH and the LHL pattern for the “spread” while other clips had significant difference between them. The “spread” sequence had a different performance and one possible reason is the video sequence having more people (more heads) than the other sequences, which points to other possible class of videos considering content type/feature. These observations suggest that specific content types may be more suited to particular scaling dimensions and that the content type might play an important role in choosing the optimal scaling dimension. From Figure 4, switching quality results in deterioration of QoE compared with zero switching (constant streaming at highest quality level). Furthermore, some researchers found different combinations of frame-rate and bitrate acceptable for certain content types [32, 33], which suggests that frame-rate scaling might be suitable for certain content types.

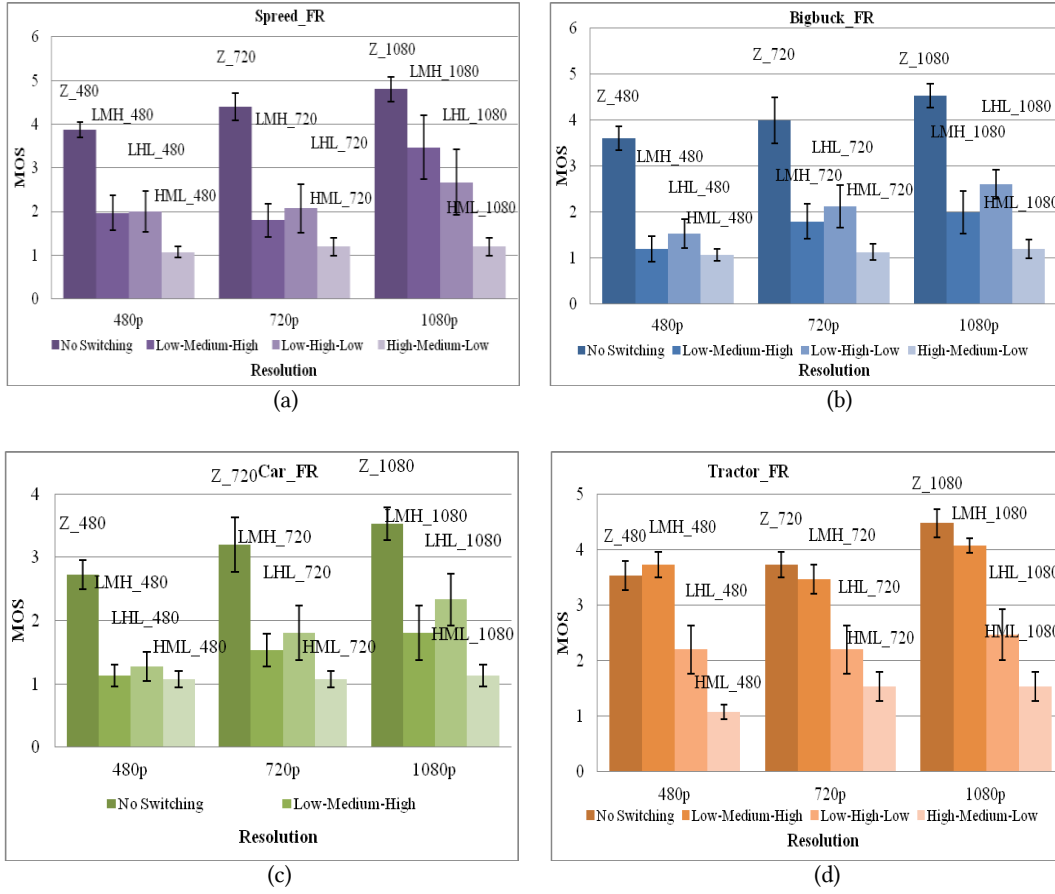


Fig. 3. MOS for frame rate scaling: (a) Spread clip, (b) Bigbuck clip, (c) Car clip, (d) Tractor clip.

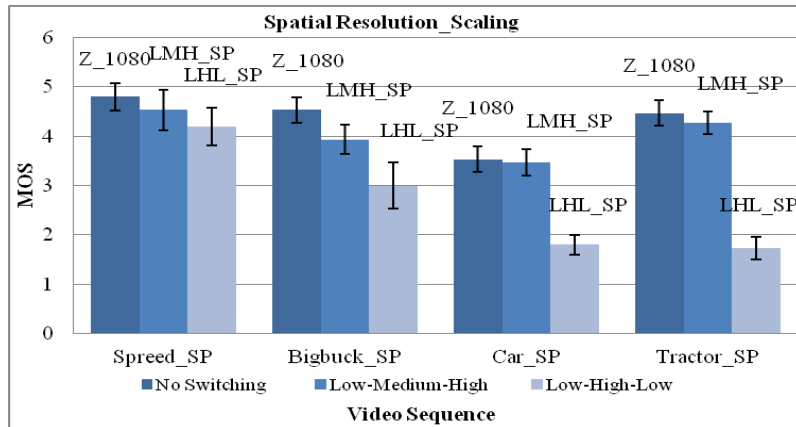
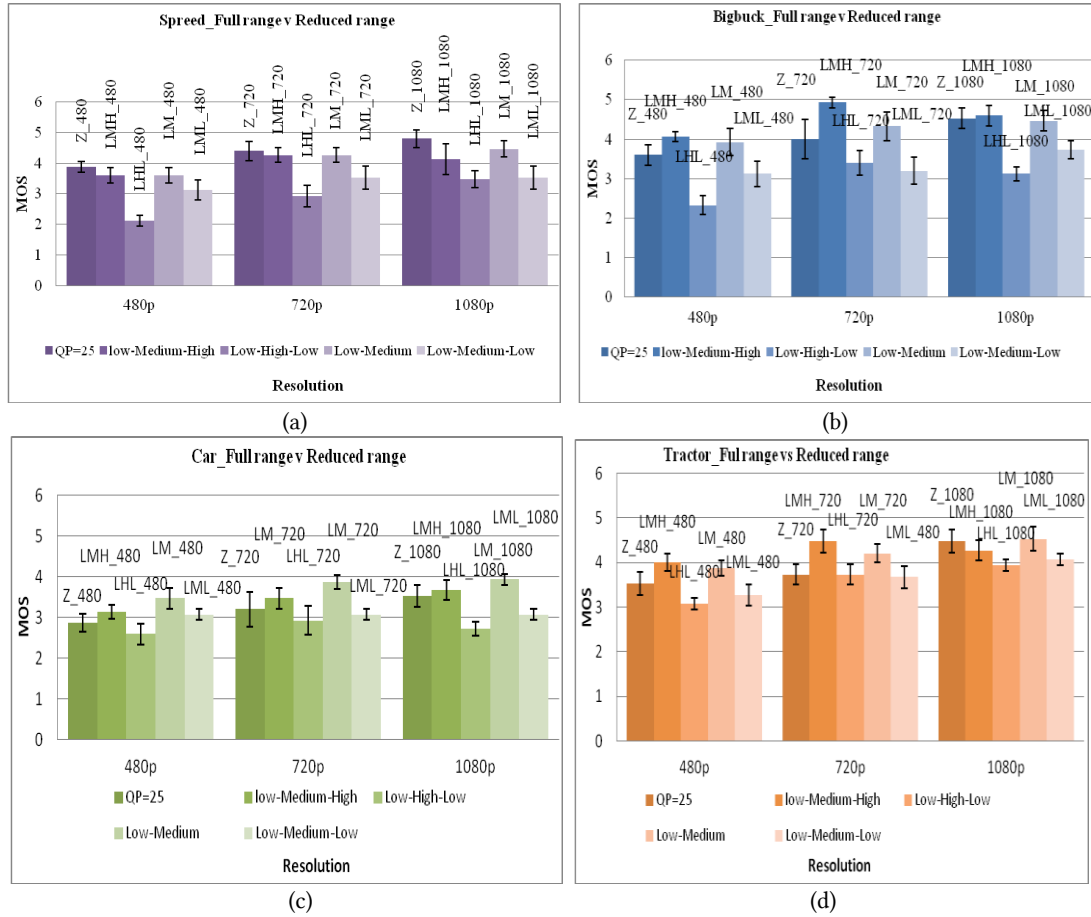


Fig. 4. MOS for spatial resolution scaling for “Spread”, “Bigbuck”, “Car” and “Tractor” clips.

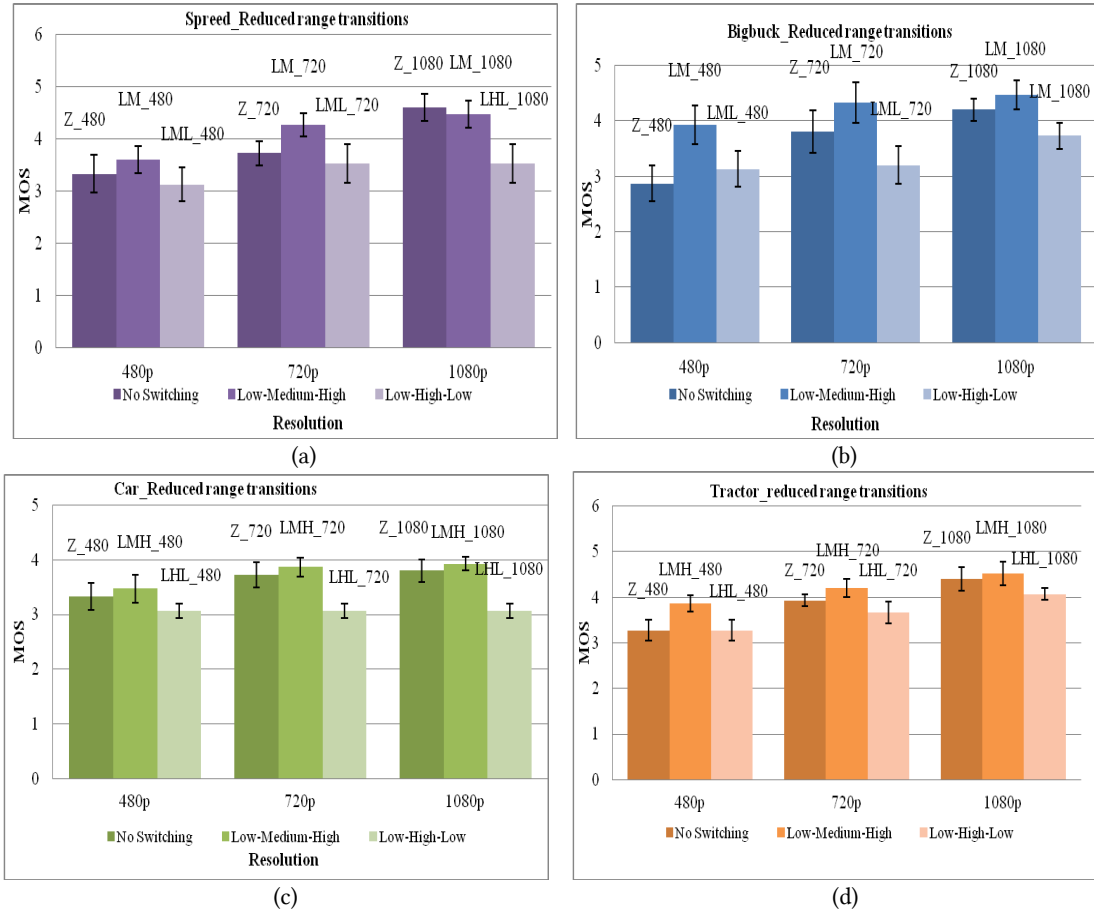


**Fig. 5. MOS comparison for frame quality scaling (QP-based) using full range transitions (involving high quality) and reduced ones (involving up to medium quality): (a) Speed clip, (b) Bigbuck clip, (c) Car clip, (d) Tractor clip.**

Even though the compression difficulty is directly related to SI/TI of a sequence [26], the results from the experiments show that the performance of different video contents in DASH cannot fully be explained with SI/TI values. The explanation will need a combination of the content type, quality scaling method, segment size, network parameters, other content features etc.

### 3.3 Further Evaluation and Hypothesis for Novel Enhancements

The experiments in the previous section involved switching quality among three dimensions. We showed that switching up quality in does not always necessarily lead to an improvement in QoE (Figure 2, 2 (d) 1080p, Figures 3 and Figure 4). We also demonstrated that subjects did not notice the expected increase in frame quality when increasing quality from the lowest quality to the highest quality and back to the lowest quality, which can be attributed to the recency effect [34]. These observations suggest that under certain conditions bandwidth can be saved while maintaining a comparable level of quality. Overall, the frame quality scaling and the spatial resolution scaling dimensions performed better (there was less variation in QoE when switching) than the frame-rate scaling dimension and will be the focus of our proposed bandwidth saving enhancements.



**Fig. 6. MOS values for frame quality scaling (QP-based) using reduced range transitions (involving up to medium quality): (a) Spread clip, (b) Bigbuck clip, (c) Car clip, (d) Tractor clip.**

In order to investigate the effect of switching to an intermediate quality level instead of the highest quality level, we repeated the experiment in the previous section by eliminating the highest quality level. Hence, we switch between two quality levels now, referred to as low and medium and by using the frame quality dimensions. In this experiment 2 as described in Table 1, we used 15 subjects, 4 video sequences (“Spread”, “Bigbuck”, “Car” and “Tractor”) as shown in Table 3, ACR scale, and the same test environment as used in experiment 1.

Figure 5 shows a comparison of the MOS levels for quality switching involving the frame quality (QP index) across the three different resolutions. The following can be observed. According to the figure, Low-Medium-High (LMH) transitions yield a higher QoE in general terms (even higher than for the non-impaired videos for “bigbuck”) compared to the LML transitions, and these are generally higher than Low-Medium (LM) switching. Differences in MOS values between LMH and LM are not significant in general terms and based on Mann-Whitney U test there was no significant difference between the LM and LMH transitions. (In any case, MOS values for LM switching generally fall in the good range of the ACR scale. Low-High-Low (LHL) transitions yield poorer QoE perceptions when compared to Low-Medium-Low (LML). This impairment in QoE suggests that the test subjects were happier with switching down from medium quality to low quality than switching from high quality down to low quality. This observation will further be investigated, as the results suggest that the amplitude of switching in negative polarity plays an important role in QoE.

In our final experiment, we set up the intermediate QP level of 30 with maximum frame rate as the baseline for streaming scenarios (frame-quality switching dimension), and evaluated the QoE perceptions when LML and LM switching was applied. Figure 6 presents the MOS results for this experiment. To be coherent with the previous section, we have referred to these scenarios as Z\_X, LM\_X, and LML\_X, where X can be one of the following resolutions (480p, 720p, 1080p). From the results in figure 6, we can conclude the following: There is a general improvement in QoE for LM switching as compared to the baseline. This is more relevant for 480p and 720p resolutions. A comparison with Figure 2 shows, that the degradation in QoE experienced by the subjects in Low to Low scenarios is more noticeable when a high quality scenario is involved (see figure 2). A possible reason for this is the amplitude of degrading when switching down from a higher quality is more noticeable to the test subjects than switching down from an intermediate quality. The results show that the higher the amplitude of switching in improving quality the higher the QoE but the higher the amplitude of decreasing quality the lower the QoE. This trend will also be investigated further with higher resolutions. The results from our experiments suggest that reducing the amplitude of switching down quality in error prone networks will improve the QoE and also increasing the bitrate does not necessarily improve the QoE. All these observations and other factors are taken into account to propose a QoE-Aware Cost-Efficient algorithm in the next section.

## 4 QoE-AWARE COST-EFFICIENT FRAMEWORK (QACEF)

### 4.1 Design Principles of the Framework

We firstly contextualise and outline our design principles. In DASH, different quality levels of videos are usually segmented into video segments of fixed duration, which allows for quality switching at Group of Pictures (GoP) boundaries. We evaluated video segments of 2s and 10s duration, being the recommendations by Microsoft and Apple, and referred henceforth as short and long segment sizes, respectively. There is a trade-off between short and long segment sizes. Short segment sizes allow finer grained adaption since the switching decision is taken faster whereas long segment sizes allow better compression and a smaller number of switches. The correct selection of the segment size will depend on the level of variation of the network conditions. With the need for optimal selection of segment size and to propose a novel adaptation algorithm after our previous experiments and initial hypothesis, we will further answer the following research questions:

- Can we combine short and long segment durations thereby achieving their combined merits?
- Can the amplitude of switching down quality be minimized to improve QoE?
- Can we achieve the recommended best practices for maintaining adequate QoE as recommended in [30]? These include avoiding multiple short quality adaptations, i.e. adaptation algorithms should not always immediately request the highest quality level if resources allow, adaptation algorithms switching to an intermediate level before switching to the highest quality level, and algorithms preventing video Stallings.

Hence, we propose combining the advantages of short and long segment sizes by using a physical short segment size of 2 seconds, which allows for faster adaptation, and a logical long segment size of 10s to minimise number of switches while also saving bandwidth. This contribution proposes a heuristic-driven QoE-aware cost-efficient framework that does not switch quickly to the highest bitrate when the network condition allows for, but rather drives the video adaption to stay temporarily at an intermediate quality level while achieving a more constant quality.

The proposed scheme makes the following assumptions: i) the duration of the segment is 2s as specified by Microsoft [6] and ii) 5 throughput measurements will yield 10s duration after a drop in network condition as recommended by Apple [7]. Based on these foundations, the proposed framework uses 2s segment sizes to achieve the same period of switching with the 10s duration of the Apple proposal, while reducing the number of potential switches, minimizing non-optimal switching, minimizing the amplitude of switching and finally allowing the video sequence to remain at an intermediate quality level when there is a presence of unstable network conditions.



The advantage of the proposed framework is fourfold. i) The increased period of switching minimizes unnecessary switching which causes flicker in unstable network environments thereby improving QoE [22]. ii) It also saves bitrate since an intermediate bitrate is temporarily chosen as the optimal bitrate that achieves comparable perceptual quality as the highest bitrate. 3) Temporarily staying at an intermediate quality level reduces the amplitude of switching down suddenly in unstable network environments. 4) The technique potentially reduces inaccurate bandwidth measurements caused by using short segment sizes [35] where short time network capacity measurements lead to non-optimum quality adaptations. Staying at an intermediate quality level achieves bandwidth savings, while also reducing power consumption. This can be beneficial to mobile devices, where battery resource is an on-going issue coupled with limited processing capacity [36]. At a glance, the proposed algorithm achieves the first two (2/3) recommended best practices for maintaining adequate QoE, as recommended by the authors in [30].

To implement this framework, an additional module has been deployed in the DASH video client. This module performs QoE analysis before recommending a video layer to be reproduced. This module is optional as it is recommended for further optimization of the system. The inclusion of this framework in the diagram aids in highlighting one of the most important components of this work. The algorithm proposed can only switch to the highest quality video layer when; a) the highest quality video layer does not exceed the available bandwidth of the network and at the same time b) only after 5 positive continuous measurements of network throughput. The algorithm uses this heuristic approach to isolate the short-term throughput variations that may happen in unstable network conditions because of the TCP congestion control mechanisms of the TCP/IP stack.

Figure 7 shows all the steps involved in the proposed framework to achieve a cost-effective QoE-aware framework for dynamic video adaptation. The video sequence is encoded into various video quality layers by the QoE-aware encoder. The intermediate quality level is determined by using the data of the previous subjective testing carried out in order to determine the optimal encoding parameters to maximize bandwidth saving whilst maximizing QoE as presented in the previous section. The aim is to achieve an intermediate layer comparable in terms of level of quality with the highest bitrate level. The encoded videos are then segmented into 2s chunks and then stored on the standard web server for delivery upon request from the DASH client. The key component of our proposed system is the DASH client, which is aware when the recommended intermediate quality level is available in the video stream. The main aim of our proposed QoE-Aware cost-efficient system is to optimize the QoE while saving substantially in network bandwidth.

The reference client (MP4Client [37]) measures the previous download rate while adapting the video streaming session according to the user's network condition. The MP4Client version used here adopts an aggressive switching approach where it switches to the highest quality level if the network permits at the first download rate measurement. Our proposed algorithm also measures the previous download rate while adapting the video streaming session but replaces the aggressive switching approach with a QoE informed approach. The QoE-Aware cost-efficient algorithm is an adaptive video streaming algorithm that selects the next optimal video layer to be streamed based on throughput measurement of the current video segment. As soon as the web server gets an http request from the DASH player, it sends an MPD file, which describes the various video layers stored on the web server. The DASH client runs the proposed algorithm to allow the decision of the video layer to be downloaded next. After each video segment is downloaded, the throughput of the current segment is measured.

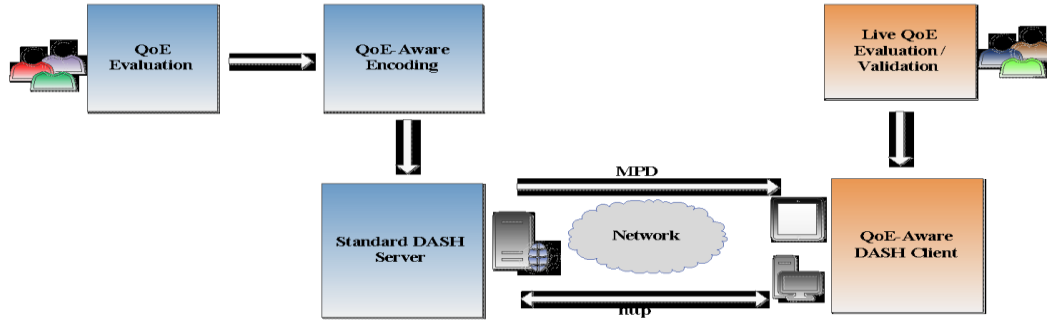


Fig. 7. Overview of QoE-aware cost-efficient system.

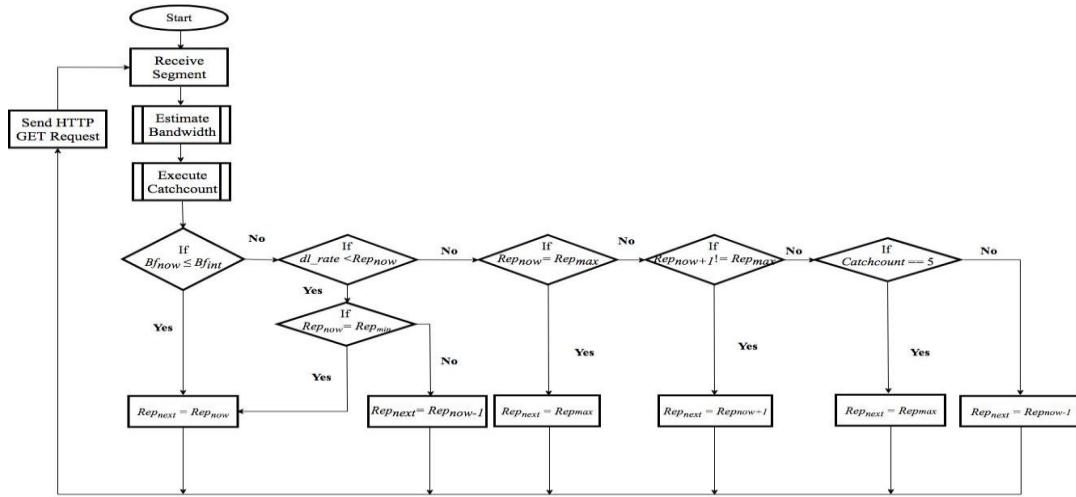


Fig. 8. Proposed QoE-aware cost-efficient algorithm for switching operations flow chart.

**Case 1.** If the bitrate of the current video layer  $Rep_{now}$  is less than the download rate, we increment a key component of our algorithm, called *Catchcount*, which is a counter that monitors positive throughput measurements. If the current buffer level  $Bf_{now}$  is less than the initial buffer threshold  $Bf_{int}$  then the minimum video layer  $Rep_{min}$  is chosen as the next segment. This case is strictly to trigger catchcount when there is a drop in network condition and also to get the video running at the minimum bitrate to allow buffer build up.

**Case 2.** If the network bandwidth available is greater or equal to the bitrate of the current video layer then If the current video layer is the highest bitrate representation  $Rep_{max}$ , then the next layer will be the same. However, if the current layer is not the highest quality layer and if the immediate next quality video level is not the highest video layer then that one is selected. However, when the next video layer is equal to the highest quality layer, then only and only if the *Catchcount* is equal to 5 that highest layer will be selected. If the *Catchcount* is less than 5 then the next to highest video quality layer will be selected instead. This layer is referred as the intermediate video quality level.

**Table 6. Streaming Scenarios for Validation**

Sequence ID	Network Impairment	Description of Scenario
BW1	(Constant) Bandwidth	This scenario describes an intermittent switching between 5s with a constant bandwidth limitation at 600KBit/s followed by 4s with no bandwidth limitation.
BW2	(Linear) Bandwidth	This scenario describes an intermittent switching between 5s with an increasing bandwidth limitation ranging from 400 to 600KBit/s in steps of 50Kbit/s, followed by 4s with no bandwidth limitation.
PL1	(Constant High) Packet Loss	This scenario describes an intermittent switching between 5s with a constant packet loss at 15% followed by 5s with no packet loss.
PL2	(Variable High) Packet Loss	This scenario describes an intermittent switching between 5s with a variable packet loss at $15\% \pm 5\%$ followed by 5s with no packet loss.
PL3	(Linear) Packet Loss	This scenario describes an intermittent switching between 5s with an increasing packet loss ranging from 5% to 35% in steps of 5% followed by 5s with no packet loss.
PL4	(Constant Medium) Packet Loss	This scenario describes an intermittent switching between 2s with a constant packet loss at 10% followed by 10s with no packet loss.
D1	(Constant Medium) Delay	This scenario describes an intermittent switching between 5s with a constant delay 300ms followed by 4s with no delay.
D2	(Constant Low) Delay	This scenario describes an intermittent switching between 2s with a constant delay 300ms followed by 8s with no delay.
D3	(Delay variation Low) Delay	This scenario describes an intermittent switching between 5s with variable delay $250\text{ms} \pm 50\text{ms}$ followed by 4s with no delay.
D4	(Linear) Delay	This scenario describes an intermittent switching between 5s with an increasing delay limitation ranging from 100 to 350ms in steps of 50ms, followed by 4s with no delay limitation.

**Case 3.** If the network bandwidth available is less than the current video layer, then if the current video layer is the lowest one then it will be selected. However, if the current one is not lowest one, then the immediately lower in quality will be selected.

**Case 4.** Finally, if the bitrate of the next video layer to be selected by the algorithm is lower than the bitrate of the current video layer then the *Catchcount* is reset to zero.

It is noted that the *Catchcount* indicates degradation in network conditions. Consequently, the algorithm will wait for 5 positive throughput measurements before selecting a segment from the highest bitrate level as the next segment  $Rep_{next}$ . For clarity, the experiments and the description of the algorithm deal only with 3 bitrate levels. However, the proposed scheme is easily adaptable to any number of video layers. This algorithm (Figure 8) can easily be adapted for  $n$  layers; we only used 3 layers in order to demonstrate how the algorithm works at a glance.

## 4.2 Evaluation Methodology and Validation

**Table 7. Video Sequences for Bandwidth Savings and QoE Evaluation**

Video Clip	Frame Rate	QP	Spatial Resolution	SI	TI	Description
Elephant Dreams	24	25, 30, 35	1280x720	35	9	Animated science fiction, high motion with two main characters exploring the twisted and dark complex wires, cogs and gears of a giant machine [40].
Tomsk2	24	25, 30, 35	1280x720	77	7	Documentary, average motion, people driving around a city and some scenes with people walking in a park [41].
Building	24	25, 30, 35	1280x720	39	6	Amateur filming, no motion and no scene changes, a moving camera filming a very large building.
Spacecraft	24	25, 30, 35	1280x720	82	10	Documentary, high motion, rapid scene changes, a space craft being launched into space.

To validate our proposed algorithm in terms of bandwidth optimization, we calculated the total bitrate streamed using our proposed QACEF algorithm and compared it against the state of the art GPAC MP4Client [37]. The comparison has been made in terms of the bandwidth savings. All the results presented in this section consider the bitrate saved as the relative subtraction between the average bitrate sent from GPAC MP4Client and the average bitrate sent from the proposed scheme.

To validate our proposed algorithm in terms of QoE, a subjective evaluation was carried out using the ACR method previously described in Table 2. The Mean Opinion Scores (MOS) was calculated and analyzed to compare the performance between the two algorithms. The steps involved are further summarized in the following subsections.

#### **4.2.1 Test Material, Conditions and Environment.**

In order to calculate the bandwidth savings and to obtain the MOS, a testbed was deployed to carry out the experiment. It is composed by 3 computers, a DASH server, a DASH client and a computer in-between acting as a networking device. This networking device has been used as a NETEM box to emulate 10 different network conditions between client and server. These conditions are summarized in Table 6. To understand the performance of the system, we adopted and modified the scenarios used by Martin et al. [38] and some parameters used by Huysegems et al. [39]. To be specific, these 10 scenarios are used to investigate the impact of different network impairments like fluctuation in bandwidth (scenarios 1-2), packet loss rate (scenarios 3-6) and network delay (scenarios 7-10). A preliminary test was carried out before defining the scenarios to determine upper and lower bounds for the impairments, considering the type of codec, video sequence and encoding parameters, among others. 4 different video sequences of 60s duration with different levels of detail (SI and TI) were chosen for this evaluation as further described in Table 7. The videos were encoded into 3 different QP values in order to create 3 different quality levels. All 10 scenarios were used to evaluate bandwidth savings using only the “elephant dream” sequence; however, only the ideal streaming (zero switching) at 25QP (24fps, 1280x720p), BW1 and D1 were selected for QoE evaluation. This is due to the difficulty of carrying out subjective evaluation, since gathering a good number of volunteers (users) can be difficult. For each of these 3 networking scenarios, 4 video sequences were used namely, “elephant dream”, “tomsk2”, “building” and “spacecraft”. Video sequences were presented to the test subjects one at a time and then the subjects were asked to rate them after

each streaming session. All the subjects were first trained on how to use the ACR scale and rate the video streaming sessions accordingly.

A total of 15 test subjects participated in the subjective testing. All the subjects were checked for normal visual acuity and had no reported health condition. The subjective testing was carried out in a controlled environment using a 22-inch TV monitor in accordance with ITU [26] guidelines.

#### **4.2.2 Final Results and Analysis.**

It should be noted that ideal streaming in our analysis and graph refers to streaming at the highest quality level without switching (zero switching).

##### **4.2.2.1 Bandwidth saving evaluation**

Figure 8(a) shows the average bandwidth savings obtained with the proposed scheme QACEF with respect to the state of the art algorithm implemented in the GPAC MP4Client. It is worth highlighting that the proposed algorithm over-performed for all the analyzed experiments, achieving improvements in some cases up to 41% in saved bandwidth. The figure shows more significant savings when the conditions of the network are subject to bandwidth (BW) or delay (D) constraints. In the case of packet loss (PL) scenarios, QACEF still performed better than the state of the art although the improvement is lower (in the range of 10%).

Figure 9(b) shows the distribution of bitrate savings for all the analyzed videos. It is noted that the achieved bandwidth savings are greater than or equal to 30% in 40% (the last two bars combined together) of the analyzed scenarios, and that overall 80% of the investigated scenarios achieved a minimum of 10% bitrate savings using the proposed scheme. Thus, our proposal QACEF can provide a significant improvement in the adaptive video streaming technology, especially considering limited bandwidth scenarios and the unpredictable nature of the future Internet. Additionally, the proposed architecture is also applicable in network environments with high bandwidth fluctuations such as mobile networks.

##### **4.2.2.2 Subjective evaluation**

It is noted that the design goal is to achieve efficiency (bandwidth saving) without compromising QoE. To this end, the results of the subjective evaluation for the proposed QACEF scheme are collected in another subjective experiment and shown in Figure 10. The QoE obtained from QACEF is compared with the QoE of the state of the art DASH Client MP4Client. Under stable network conditions, the proposed algorithm achieved a MOS value of 4.73, comparable with the reference algorithm 4.53. The error bars for 95% confidence interval do overlap, which suggests that there may or may not be any statistical significance between the two algorithms. Therefore, the results show comparable QoE for the end users. The results indicate that the end users were satisfied with the gradual change in quality levels, and used to a particular level of quality before moving to the next level of quality.

When the network conditions is unstable in terms of bandwidth at 600Kbits/s for the elephants dream video sequence, the average MOS scores obtained for both algorithms were 1.7, which shows that the test subjects did not see any perceptual quality difference between them. However, it is worth recalling that our algorithm achieved a significant bitrate saving compared with the reference algorithm. With an unstable network in terms of delay at 300ms, our proposed algorithm has a MOS value of 3.73 while the reference algorithm had a MOS value of 3.86. We show that they are also comparable under these circumstances while our proposal still provides significant savings in terms of bandwidth. Using the Mann-Whitney U test, there was no significant difference between our proposed algorithm and the reference algorithm for the “elephantdream” clip, which is an animation video with lower SI/TI. For the “tomsk2” clip, users generally had a slightly better average QoE with our algorithm especially at unstable network conditions (BW1) there was a more significant difference between our proposed algorithm and the reference DASH client algorithm. For the other sequences, our algorithms consistently performed slightly better in terms of average QoE (and thus comparable QoE) in all the network conditions analyzed whilst gaining a significant efficiency improvement in terms of the bandwidth saved. These results show a step forward in advancing the current state of the art.

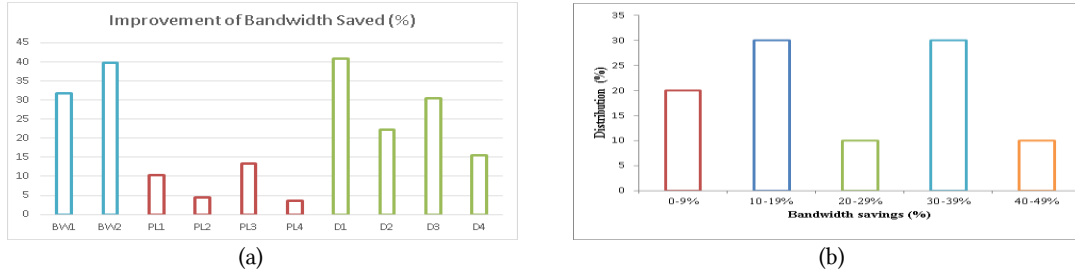


Fig. 9. (a) Bandwidth savings for all scenarios in one glance. (b) Distribution of bandwidth savings.

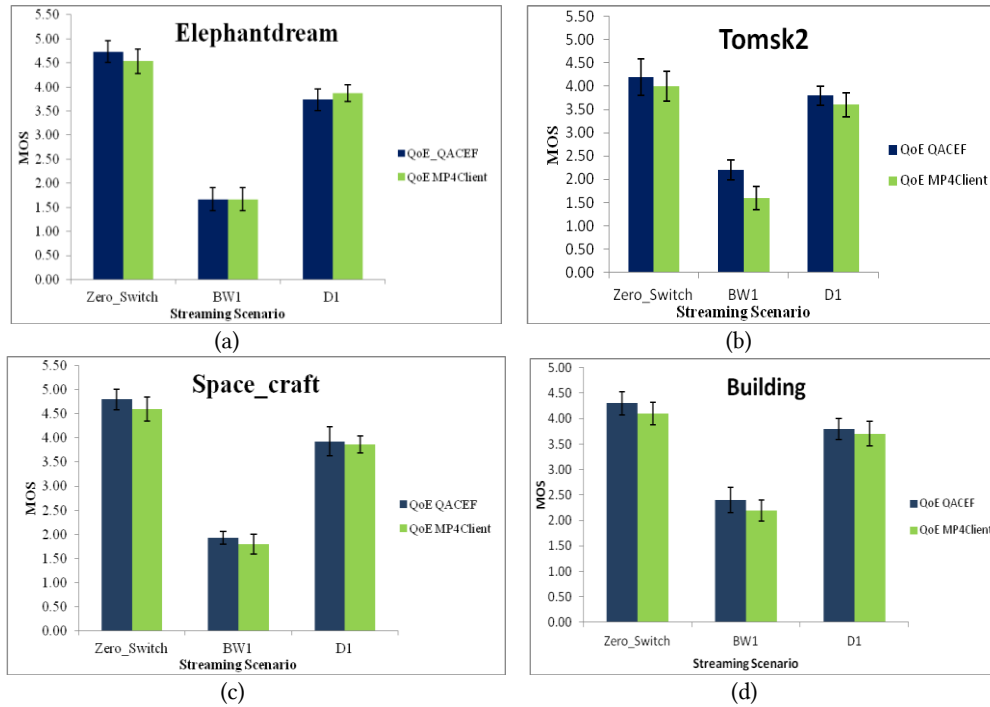


Fig. 10. MOS for different streaming scenarios and algorithm: (a) Elephants dream clip, (b) Tomsk2 clip, (c) Spacecraft clip, (d) Building clip.

## 5 CONCLUSIONS

In this paper, we have empirically studied the perceived impact of different adaptive video streaming quality scaling dimensions in various streaming scenarios. This work was motivated by the lack of evidence on the interaction of the latest coding standard with different quality scaling dimensions and the need to gain deeper understanding and more evidence into some of the visual artefacts introduced by the different quality scaling dimensions. Informed by this empirical study, we have defined optimal operating conditions and proposed novel techniques to improve the performance of the adaptive streaming technique. Specifically, with a forward-looking and future-proof vision for the next-generation video streaming applications, we have employed the H.265/HEVC standard to investigate different switching patterns and different quality scaling dimensions. We have compared our results with some contradicting results obtained in the literatures using the H.264/SVC or H.264/AVC standards while verifying the robustness of some other claims. The streaming scenarios have been

carefully chosen to demonstrate basic stepwise quality switching operations in everyday video streaming with a focus on selected factors such as the complexity within the video sequence, an innovative interaction between quality scaling dimensions, and the polarity of switching.

We initially ran some experiments to gain insights into factors determining perceptual quality in adaptive video streaming in H.265/HEVC. We have observed significant interplay among the spatial, temporal and quality scaling dimensions. Our analysis and results demonstrate that going up a bitrate level does not necessarily mean improving the QoE in certain streaming scenarios. In these scenarios, we have found that staying at an intermediate bitrate level aids in achieving bitrate savings whilst achieving comparable quality to the highest level of quality. We have confirmed the significant impact of switching polarity on the QoE and how reducing the amplitude of switching in unstable network environments to optimize QoE. The results also highlight the relationship between switching frequency and polarity of switching. Furthermore, the influence of the complexity within the video sequence has been established to have a significant influence on the different quality scaling dimensions. Finally, following our results and analysis, we have proposed a novel adaptation scheme that saves up to 41% bitrate while achieving a comparable level of quality. In future work, we plan to further investigate the link between the performance for different types of videos and their characteristics, for example, their content types, spatial and temporal features and so on.

## REFERENCES

- [1] Cisco VNI: Forecast and Methodology, 2015-2020 <http://www.cisco.com/c/en/us/solutions/collateral//service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>.
- [2] ITU-T, Recommendation H.265 (V4), High efficiency video coding, Dec. 2016, [online] Available: <http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=11885>.
- [3] NGMN Alliance, "Perspectives on Vertical Industries and Implications for 5G", White Paper, Jun. 2016, [online] Available: [https://www.ngmn.org/uploads/media/160610\\_NGMN\\_Perspectives\\_on\\_Vertical\\_Industries\\_and\\_Implications\\_for\\_5G\\_v1\\_0.pdf](https://www.ngmn.org/uploads/media/160610_NGMN_Perspectives_on_Vertical_Industries_and_Implications_for_5G_v1_0.pdf).
- [4] Iraj Sodagar. 2011. The MPEG-DASH Standard for Multimedia Streaming Over the Internet. *IEEE Multimedia* 18, 4 (2011), 62–67. DOI:<http://dx.doi.org/10.1109/mmul.2011.71>
- [5] Gaetano Carlucci, Luca De Cicco, and Saverio Mascolo. HTTP over UDP: an Experimental Investigation of QUIC, *Proceedings of the 30<sup>th</sup> Annual ACM Symposium on Applied Computing*, ACM 2015.
- [6] Alex Zambelli. 2009. Microsoft Smooth Streaming. IIS Smooth Streaming Technical Overview. From [http://www.bogotobogo.com/VideoStreaming/Files/iis8/IIS\\_Smooth\\_Streaming\\_Technical\\_Overview.pdf](http://www.bogotobogo.com/VideoStreaming/Files/iis8/IIS_Smooth_Streaming_Technical_Overview.pdf). (Last access: June 2017).
- [7] Roger Pantos. 2017. HTTP Live Streaming. IETF draft (May 2017) from <https://tools.ietf.org/html/draft-pantos-http-live-streaming-23> (last access: June 2017).
- [8] Saamer Akshabi, Sethumadhavan Narayanaswamy, Ali C. Begen, and Constantine Dovrolis. 2012. An experimental evaluation of rate-adaptive video players over HTTP. *Signal Processing: Image Communication* 27, 4 (2012), 271–287.
- [9] Zhi Li, Xiaoqing Zhu, Joshua Gahm, Rong Pan, Hao Hu, Ali C. Begen, and David Oran. 2014. Probe and Adapt: Rate Adaptation for HTTP Video Streaming At Scale. *IEEE Journal on Selected Areas in Communications* 32, 4 (2014), 719–733.
- [10] Te-Yuan Huang, Ramesh Johari, Nick Mckeown, Matthew Trunnell, and Mark Watson. 2015. A buffer-based approach to rate adaptation: Evidence from a large video streaming service. *ACM SIGCOMM Computer Communication Review* 44.4 (2015): 187–198.
- [11] Eren Gürses, Gozde Bozdagi Akar, and Nail Akar. 2005. A simple and effective mechanism for stored video streaming with TCP transport and server-side adaptive frame discards. *Computer Networks* 48, 4 (2005), 489–501.
- [12] Jeroen Van Der Hooft, Stefano Petrangeli, Tim Wauters, Rafael Huysegems, Patrice Rondao Alfai, Tom Bostoen, and Filip De Turck. 2016. HTTP/2-Based Adaptive Streaming of HEVC Video Over 4G/LTE Networks. *IEEE Communications Letters* 20, 11 (2016), 2177–2180.
- [13] Nicola Cranley, Philip Perry, and Liam Murphy. 2006. User perception of adapting video quality. *International Journal of Human-Computer Studies* 64, 8 (2006), 637–647.
- [14] Michael Zink, Jens Schmitt, and Ralf Steinmetz. 2005. Layer-encoded video in scalable adaptive streaming. *IEEE Transactions on Multimedia* 7, 1 (2005), 75–84.
- [15] Liu Yitong, Shen Yun, Mao Yinian, Liu Jing, Lin Qi, and Yang Dacheng. 2013. A study on Quality of Experience for adaptive streaming service. 2013 *IEEE International Conference on Communications Workshops (ICC)* (2013).
- [16] Michael Grafl and Christian Timmerer. 2013. Representation Switch Smoothing for Adaptive HTTP Streaming. 4th *International Workshop on Perceptual Quality of Systems (PQS)* (February 2013).
- [17] Demóstenes Z. Rodriguez, Zhou Wang, Renata L. Rosa, and Graça Bressan. 2014. The impact of video-quality-level switching on user quality of experience in dynamic adaptive streaming over HTTP. *EURASIP Journal on Wireless Communications and Networking* 2014, 1 (2014).
- [18] David C. Robinson, Yves Jutras, and Viorel Craciun. 2012. Subjective Video Quality Assessment of HTTP Adaptive Streaming Technologies. *Bell Labs Technical Journal* 16, 4 (2012), 5–23.
- [19] Tobias Hossfeld, Sebastian Egger, Raimund Schatz, Markus Fiedler, Kathrin Masuch, and Charlott Lorentzen. 2012. Initial delay vs. interruptions: Between the devil and the deep blue sea. 2012 *Fourth International Workshop on Quality of Multimedia Experience* (2012).
- [20] S. Shunmuga Krishnan and Ramesh K. Sitaraman. 2013. Video Stream Quality Impacts Viewer Behavior: Inferring Causality Using Quasi-Experimental Designs. *IEEE/ACM Transactions on Networking* 21, 6 (2013), 2001–2014.
- [21] Quan Huynh-Thu and Mohammed Ghanbari. 2006. Impact of jitter and jerkiness on perceived video quality. In *Proc. Workshop on Video Processing and Quality Metrics*. 2006.
- [22] Pengpeng Ni, Ragnhild Eg, Alexander Eichhorn, Carsten Griwodz, and Pål Halvorsen. 2011. Flicker effects in adaptive video streaming to handheld devices. *ACM international conference on Multimedia - MM* 11(2011).

- [23] Samira Tavakoli, Kjell Brunnström, Kun Wang, Börje Andrén, Muhammad Shahid, and Narciso Garcia. 2014. Subjective quality assessment of an adaptive video streaming model. *Image Quality and System Performance XI* (July 2014).
- [24] Quan Huynh-Thu and M. Ghanbari. 2008. Temporal Aspect of Perceived Quality in Mobile Video Broadcasting. *IEEE Transactions on Broadcasting* 54, 3 (2008), 641–651.
- [25] Q. Huynh-Thu and M. Ghanbari. 2008. Scope of validity of PSNR in image/video quality assessment. *Electronics Letters* 44, 13 (2008), 800.
- [26] ITU. 2008. ITU-T Recommendation P.910: Subjective video quality assessment methods for multimedia applications. (2008)
- [27] ITU. 2012. ITU-R Recommendation BT 500-13: Methodology for the subjective assessment of the quality of television pictures. (2012)
- [28] Samira Tavakoli, Sebastian Egger, Michael Seufert, Raimund Schatz, KjellBrunnstrom, and Narciso Garcia. 2016. Perceptual Quality of HTTP Adaptive Streaming Strategies: Cross-Experimental Analysis of Multi-Laboratory and Crowdsourced Subjective Studies. *IEEE Journal on Selected Areas in Communications* 34, 8 (2016), 2141–2153.
- [29] Blender Institute Production. Bigbuck bunny. Retrieved December 20, 2016 from <https://peach.blender.org/download/>
- [30] Nicolas Staelens, Paulien Coppens, Niels Van Kets, Glenn Van Wallendael, Wendy Van den Broeck, Jan De Cock and Filip De Turck. On the impact of video stalling and video quality in the case of camera switching during adaptive video streaming of sports content. *Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, (2015).
- [31] Nicolas Staelens, Jonas De Meulenaere, Maxim Claeys, Glenn Van Wallendael, Wendy Van den Broeck, Jan De Cock, Rik Van de Walle, Piet Demeester and Filip De Turck. Subjective quality assessment of longer duration video sequences delivered over HTTP adaptive streaming to tablet devices. *IEEE Transactions on Broadcasting* 60.4 (2014). 707-714.
- [32] John D McCarrthy, M. Angela Sasse, and Dimitros Miras. Sharp or Smooth? Comparing the effects of quantization vs frame rate for streamed video. *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM 2004.
- [33] Stefan Winkler and Christof Faller. Maximizing audiovisual quality at low bitrates. *Proc International workshop on Video Proc and Quality Metrics for Consumer Electronics* (2005).
- [34] Tobias Hobfeld, Michael Seufert, Christian Sieber, and Thomas Zinner. Assessing effect sizes of influence factors towards a QoE model for HTTP adaptive streaming. *IEEE Sixth International Workshop on Quality of Multimedia Experience (QoMEX)* (2014).
- [35] Chenghao Liu, Imad Bouazizi, and Moncef Gabbouj. 2011. Rate adaptation for adaptive HTTP streaming. *Proceedings of the second annual ACM conference on Multimedia systems - MMSys* 11(2011).
- [36] Is-Haka Mkwawa and Lingfen Sun, Power-driven VoIP quality adaptation over WLAN in mobile devices. *IEEE Globecom Workshops (GC Workshops)*, (2012), 1276-1281.
- [37] GPAC Multimedia Project. GPAC MP4Client. Retrieved January 25, 2013 from <https://gpac.wp.imt.fr/downloads>
- [38] Jim Martin, Yunhui Fu, Nicholas Wourms, and Terry Shaw. 2013. Characterizing Netflix bandwidth consumption. 2013 IEEE 10th Consumer Communications and Networking Conference (CCNC) (2013). DOI:<http://dx.doi.org/10.1109/ccnc.2013.6488451>
- [39] Rafael Huysegems, Jeroen Van Der Hoof, Tom Bostoen, Patrice Rondao Alface, Stefano Petrangeli, Tim Wauters, and Filip De Turck. 2015. HTTP/2-Based Methods to Improve the Live Experience of Adaptive Streaming. *Proceedings of the 23rd ACM international conference on Multimedia - MM* 15(2015).
- [40] Blender Foundation. Elephant dream. Retrieved December 20, 2016 from <https://orange.blender.org/>
- [41] ElecCard Video Compression guru. Toms part 2. Retrieved February 21, 2017 from <https://www.elecCard.com/videos/>